PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 5 :
C12N 15/29, 15/74, A01N 63/00, C12N 1/21, A01H 5/00, 5/10 — A1

(11) International Publication Number: WO 94/16076

(43) International Publication Date: 21 July 1994 (21.07.94)

(21) International Application Number: PCT/GB94/00012

(22) International Filing Date: 5 January 1994 (05.01.94)

(30) Priority Data:
9300281.4 8 January 1993 (08.01.93) GB

(71) Applicant *(for all designated States except US)*: ZENECA LIMITED [GB/GB]; Imperial Chemical House, 9 Millbank, London SW1P 3JF (GB).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: DUBOCK, Adrian, Christopher [GB/GB]; Courts Farm, Fernhurst, Haslemere, Surrey GU27 3JF (GB). POWELL, Keith, Adrian [GB/GB]; 6 Pitts Close, Emmetts Park, Binfield, Nr. Bracknell, Berkshire RG12 5ES (GB). REES, Sarah, Bronwen [GB/GB]; 32 Micheldever Way, Forest Park, Bracknell, Berkshire RG12 3XX (GB).

(74) Agent: ROBERTS, Timothy, Wace; Group Patent Services Dept., Imperial Chemical Industries Plc, P.O. Box 6, Bessemer Road, Welwyn Garden City, Herts AL7 1HD (GB).

(81) Designated States: AU, BB, BG, BR, BY, CA, CZ, FI, GB, HU, JP, KP, KR, KZ, LK, MG, MN, MW, NO, NZ, PL, RO, RU, SD, SK, UA, US, VN, European patent (AT, BE, CH, DE, DK, ES, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG).

Published
With international search report.

(54) Title: ANTIMICROBIAL-PROTEIN-PRODUCING ENDOSYMBIOTIC MICROORGANISMS

(57) Abstract

There is provided a method of producing an antimicrobial-protein-producing microorganism capable of entering into an endosymbiotic relationship with a plant host comprising the combination of genetic material encoding a plant-derived antimicrobial protein with an endophyte. Examples of potent plant-derived antimicrobial proteins are given. A method for protecting a plant host from disease comprising treating the plant host with the antimicrobial-protein-producing microorganism is described.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AT	Austria	**GB**	United Kingdom	**MR**	Mauritania
AU	Australia	**GE**	Georgia	**MW**	Malawi
BB	Barbados	**GN**	Guinea	**NE**	Niger
BE	Belgium	**GR**	Greece	**NL**	Netherlands
BF	Burkina Faso	**HU**	Hungary	**NO**	Norway
BG	Bulgaria	**IE**	Ireland	**NZ**	New Zealand
BJ	Benin	**IT**	Italy	**PL**	Poland
BR	Brazil	**JP**	Japan	**PT**	Portugal
BY	Belarus	**KE**	Kenya	**RO**	Romania
CA	Canada	**KG**	Kyrgystan	**RU**	Russian Federation
CF	Central African Republic	**KP**	Democratic People's Republic of Korea	**SD**	Sudan
CG	Congo			**SE**	Sweden
CH	Switzerland	**KR**	Republic of Korea	**SI**	Slovenia
CI	Côte d'Ivoire	**KZ**	Kazakhstan	**SK**	Slovakia
CM	Cameroon	**LI**	Liechtenstein	**SN**	Senegal
CN	China	**LK**	Sri Lanka	**TD**	Chad
CS	Czechoslovakia	**LU**	Luxembourg	**TG**	Togo
CZ	Czech Republic	**LV**	Latvia	**TJ**	Tajikistan
DE	Germany	**MC**	Monaco	**TT**	Trinidad and Tobago
DK	Denmark	**MD**	Republic of Moldova	**UA**	Ukraine
ES	Spain	**MG**	Madagascar	**US**	United States of America
FI	Finland	**ML**	Mali	**UZ**	Uzbekistan
FR	France	**MN**	Mongolia	**VN**	Viet Nam
GA	Gabon				

ANTIMICROBIAL-PROTEIN-PRODUCING ENDOSYMBIOTIC MICRO-ORGANISMS

This invention relates to endosymbiotic micro-organisms having the ability to produce plant-derived antimicrobial proteins.

In this context, 'antimicrobial' proteins are defined as proteins possessing at least one of the following activities: antifungal activity (which may include anti-yeast activity); antibacterial activity. Activity includes a range of antagonistic effects resulting in partial inhibition or death. 'Plant-derived' proteins are capable of being isolated from the seed or other parts of one or more plant species.

Various proteins with antimicrobial activity have been isolated from plant sources, and such proteins are often believed to take part in host defence mechanisms directed against invading or competing micro-organisms. Some of the proteins are well-characterised, and their amino acid sequence may be known. In some cases, the cDNA or gene encoding the protein has also been isolated and sequenced.

To keep out potential invaders, plants produce a wide array of antifungal compounds, either in a constitutive or an inducible manner. Several classes of proteins with antifungal properties have now been identified, including chitinases, beta-1,3-glucanases, ribosome-inactivating proteins, thionins, chitin-binding lectins and zeamatins. These proteins have gained considerable attention as they could potentially be used as biocontrol agents.

The chitinases (Schlumbaum et al, 1986, Nature, 324, 363-367) and beta-1,3-glucanases have weak activities by themselves, and are only inhibitory to plant pathogens when applied in combination (Mauch et al, 1988, Plant Physiol, 88, 936-942). The chitin-binding lectins can also be classified as rather weak antifungal factors (Broekaert et al, 1989, Science, 245, 1100-1102; Van Parijs et al, 1991, Planta, 183, 258-264). Zeamatin is a more potent antifungal protein but its activity is strongly reduced by the presence of ions at physiological concentrations (Roberts and Selitnermikoff, 1990, G Gen Microbiol, 136, 2150-2155). Permatins are also known plant antifungal proteins (Vigers et al, 1991, Molec Plant-Microbe Interact, 4, 315-323; Woloshuk et al, 1991, Plant Cell, 3, 619-628). Finally, thionins (Apel et al, 1990, Physiol Plant, 80, 315-321) and ribosome-inactivating proteins (Roberts and Selitrennikoff, 1986, Biosci Rep, 6, 19-29; Leah et al, 1991, J Biol Chem, 266, 1564-1573) have antifungal activity and are known to be toxic for human cells (Carrasco et al, 1981, Eur J Biochem, 116, 185-189; Vernon et al, 1985, Arch Biochem Biophys, 238, 18-29; Stirpe and Barbieri, 1986, FEBS Lett, 195, 1-8).

Other groups of potent antimicrobial proteins with broad spectrum activity against plant pathogenic fungi (and often some antibacterial activity) are capable of isolation from certain plant species. We have previously described the structural and antifungal properties of several such proteins, including:

the small-sized cysteine-rich proteins Mj-AMP1 (antimicrobial protein 1) and Mj-AMP2 occurring in

seeds of Mirabilis jalapa (Cammue BPA et al, 1992, J Biol Chem, 267:2228-2233; International Application Publication Number WO92/15691 published on 17 September 1992);

Ac-AMP1 and Ac-AMP2 from Amaranthus caudatus seeds (Broekaert WF et al, 1992, Biochemistry, 37:4308-4314; International Application Publication Number WO92/21699 published on 10 December 1992);

Ca-AMP1 from Capsicum annuum, Bm-AMP1 from Briza maxima and related proteins found in other plants including Delphinium, Catapodium, Baptisia and Microsensis species (International Patent Application Number PCT/GB93/02179 filed on 22 October 1993);

Rs-AFP1 (antifungal protein 1) and Rs-AFP2 from seeds of Raphanus sativus (Terras FRG et al, 1992, J Biol Chem, 267:15301-13309) and related proteins such as Bn-AFP1 and Bn-AFP2 from Brassica napus, Br-AFP1 and Br-AFP2 from Brassica rapa, Sa-AFP1 and Sa-AFP2 from Sinapis alba, At-AFP1 from Arabidopsis thaliana, Dm-AMP1 and Dm-AMP2 from Dahlia merckii, Cb-AMP1 and Cb-AMP2 from Cnicus benedictus, Lc-AFP from Lathyrus cicera, Ct-AMP1 and Ct-AMP2 from Clitoria ternatea (International Patent Application Publication Number WO93/05153 published 18 March 1993);

Rs-nsLTP (non-specific lipid transfer protein) from Raphanus sativus (International Patent Application Publication Number WO93/05153 published 18 March 1993).

These publications are specifically incorporated herein by reference.

These and other plant-derived antimicrobial proteins are useful as fungicides or antibiotics to improve the disease-resistance or disease-tolerance

of crops either during the life of the plant or for post-harvest crop protection. The proteins may be extracted from plant tissue or produced by expression within micro-organisms. Exposure of a plant pathogen to an antimicrobial protein may be achieved by application of the protein to plant parts using standard agricultural techniques (eg surface spraying). The proteins may also be used to combat fungal or bacterial disease by expression within plant bodies (rather than just at the surface). DNA encoding the antimicrobial proteins (which may be a cDNA clone, a genomic DNA clone or DNA manufactured using a standard nucleic acid synthesiser) may be transformed into a plant, and the proteins expressed within transgenic plants.

It is an object of the present invention to provide an alternative method to deliver the plant-derived antimicrobial protein to its desired site of action. Such a method should be generally applicable to a wide range of plant species and may be easier or more effective than other methods.

Certain micro-organisms have the ability to enter into non-pathogenic endosymbiotic relationships with a plant host. These naturally-occurring micro-organisms, hereinafter called 'endophytes', are capable of infecting the plant host and being harboured within the plant but create no visible manifestations of disease. Such organisms include mutualistic and commensalistic endophytic organisms. The range of endophytes also includes organisms which can exist in the vascular tissues of the plant and organisms which can exist within the intercellular spaces of the plant.

A method of endophyte-enhanced protection of plants has been described in a series of patent applications by Crop Genetics International Corporation, which are discussed below and incorporated specifically herein by reference.

International Application Publication Number WO90/13224 (published 15 November 1990) describes the introduction of an endophytic bacterium into a commercially-valuable plant (such as tobacco, potato, muskmelon) to enhance protection against disease (such as tobacco mosaic virus (TMV), Pseudomonas syringae pv. tabaci, Clavibacter michiganese subsp. michiganese, potato virus X and Y, Fusarium sp. and other vascular wilt fungi). The endophyte is preferably Clavibacter xyli subsp. cynodontis (Cxc). The endophyte may be introduced into the plant by several methods including impregnating the seed with a suspension of the endophyte, using a seed coating, injecting the plant, and using a soil or foliar drench.

The endophyte may be unmodified, genetically modified (as discussed below) or formulated with other components to provide additional beneficial properties.

The endophyte may be genetically modified to produce agricultural chemicals. In this case, genetic material is derived from an agricultural-chemical-producing micro-organism and combined with a suitable endophyte. Combination of genetic material is achieved by:

(a) forming a fusion hybrid between an endophytic bacterium and an agricultural-chemical-producing bacterium (European Patent

Publication Number EP-125468-B1, published 28 October 1992); or

(b) the use of recombinant techniques (insertion of DNA encoding an agricultural chemical); for example, transforming the endophyte with an expression vector which directs production of an agricultural chemical (International Application Publication Number WO91/10363, published 24 July 1991 and International Application Publication Number WO87/03303, published 4 June 1987).

Use of the modified endophyte can improve the disease tolerance of a plant host (when compared to direct application of the agrochemical or agrochemical-producing-bacterium). The endophyte may be further improved by additional genetic modification using natural or artificial techniques (such as mutagenesis). For example, the endophyte may be modified to excrete the agricultural chemical in a particular form.

The source of DNA encoding the agricultural chemical is a suitable micro-organism. Such agricultural-chemical-producing micro-organisms are described in Table I (page 27) of International Application Publication Number WO91/10363 and include a wide variety of micro-organisms producing antibiotics, antifungal agents, antibacterial agents, antiviral agents, insecticides, nematocides, miticides, herbicides, fertilisers (nitrogen-fixing or phosphate solubilising agents), plant growth regulators or anti-feeding agents.

Suitable endophytes include Agrobacterium tumefaciens, Erwinia carotovora, Pseudomonas solanacearum, Pseudomonas syringae, Xanthomonas

campestris, Streptomyces ipomoea for dicotyledonous plants; Erwinia stewartii, Xanthomonas campestris, Azospirillum lipoferum, Azospirillum brasilense, Pseudomonas syringae for monocotyledonous plants. Clavibacter xyli subsp. xyli and Clavibacter xyli subsp. cynodontis (Cxc) are particularly useful for grasses such as maize, sorghum and the like.

The agricultural-chemical-producing endophytes may be used to enhance disease protection in any plant, including those producing fruit, vegetables and flowers, trees, field and row plants such as corn, sorghum, wheat, barley, oats, rice, brome grass, sugar cane, cotton, potatoes, tomatoes, cabbage, cauliflower, broccoli, melons, cucumbers.

International Application Publication Number WO88/09114 (published 1 December 1988) describes plants colonised by beneficial endophytic micro-organisms obtained by germination of seeds impregnated with the endophytes. The endophyte may be a strain of the genus Clavibacter or Rhizobium, and may be genetically modified to produce an agricultural chemical. The seed may be from the Gramineae, Leguminosae or Halvaceae family. International Application Publication Number WO91/11907 (published 22 August 1991) describes the production of modified seed (particularly rice) containing an unmodified or modified endophyte (particularly Cxc) to produce a plant of reduced stature.

Crop Genetics International have already developed a corn bioinsecticide based upon this endophyte technology (trademark: INCIDE Technology). The INCIDE bioinsecticide consists of the endophyte

Clavibacter xyli subsp. cynodonits (Cxc) which has been genetically modified with an endotoxin gene derived from the bacterium Bacillus thuringiensis, and thus expresses a protein which is toxic to certain insect larvae. If corn seed is inoculated with the INCIDE vaccine, the modified Cxc inhabits the vascular tissue of plants grown from this seed and the crop is protected from attack by cornborer larvae. However, there may be an associated yield reduction in certain crop species or varieties (Agrow, 13/11/92, no 172, p 6).

European Patent Application Publication Number 185005 (Monsanto Co, published 18 June 1986) also describes a "plant-colonizing micro-organism" (herein called an endophyte) which has been genetically modified to express a B thuringiensis protein.

When using an agricultural-chemical-producing endophyte to enhance disease protection in a plant, the source of DNA encoding the agricultural chemical is a suitable micro-organism. Plant-derived DNA sequences encoding antimicrobial proteins have not previously been used to modify the endophytes.

To improve disease-resistance or disease-tolerance of crops, plant-derived antimicrobial proteins may be produced within the crop plant by expression of a gene incorported into the plant genome. This may involve over-expression of an inherent protein or expression of a protein derived from another plant species. We now provide the means to express the antimicrobial protein within the crop plant without requiring plant transformation.

According to the invention, there is provided a method of producing antimicrobial-protein-producing micro-organisms capable of entering into endosymbiotic relationships with a plant host comprising the combination of genetic material encoding a plant-derived antimicrobial protein with an endophyte.

There is further provided antimicrobial-protein-producing micro-organisms produced according to the method of the invention, and seed and plants treated with said micro-organisms. Antimicrobial protein may thus be expressed within the plant by an endophyte rather than being directly expressed by the host crop plant.

As noted above, use of a genetically modified endophyte to deliver an agricultural chemical (including antifungal agents) has been described. However, the agricultural chemical was expressed from a gene derived from another micro-organism (usually a bacterium). Genes encoding plant-derived antimicrobial proteins have not been previously used (or suggested) to modify the endophyte.

Examples of plants which may be protected using the antimicrobial-protein-producing micro-organisms include field crops, cereals, fruit and vegetables such as: canola, oil seed rape, sunflower, tobacco, sugarbeet, cotton, soya, maize, wheat, barley, rice, sorghum, tomatoes, mangoes, peaches, apples, pears, strawberries, bananas, melons, potatoes, carrot, lettuce, cabbage, onion.

DNA encoding any plant-derived antimicrobial

protein may be used in the method according to the invention (for example, DNA encoding chitinases, hevein, lectins, thionins, etc).

By way of example only, DNA encoding the following plant-derived antimicrobial proteins may be used in the method according to the invention: Mj-AMP1, Mj-AMP2, Ac-AMP1, Ac-AMP2, Ca-AMP1, Bm-AMP1, Rs-AFP1, Rs-AFP2, Br-AFP1, Br-AFP2, Bn-AFP1, Bn-AFP2, Sa-AFP1, Sa-AFP2, At-AFP1, Dm-AMP1, Dm-AMP2, Cb-AMP1, Cb-AMP2, Lc-AFP, Ct-AMP1, Ct-AMP2, Rs-nsLTP. These proteins show a high level and wide spectrum of antifungal activity, and will be particularly useful for improving disease-resistance or disease-tolerance in crops. In particular, one or more of these potent antimicrobial proteins may be used in conjunction with a slower-growing endophyte as a relatively low dose of the highly active protein may be needed to provide disease protection. The presence of a slower-growing endophyte may result in less diversion of the host plant's metabolic resources, maintaining crop yield. In addition, use of these potent plant-derived antimicrobial proteins may extend the range of plant hosts most suitable as targets for this type of disease protection. Even endophytes which are relatively poor colonisers of certain plant species (such as Cxc on wheat) may be engineered to express one or more of the potent proteins to give the desired level of protection to the host plant.

The invention will now be described by way of example only, with reference to the Sequence Listing in which:

SEQ ID NO:1 is the amino acid sequence of

Mj-AMP1.

SEQ ID NO:2 is the amino acid sequence of Mj-AMP2.

SEQ ID NO:3 is the nucleotide sequence of Mj-AMP1.

SEQ ID NO:4 is the amino acid sequence of Mj-AMP1 deduced from SEQ ID NO:3.

SEQ ID NO:5 is the nucleotide sequence of Mj-AMP2.

SEQ ID NO:6 is the amino acid sequence of Mj-AMP2 deduced from SEQ ID NO:5.

SEQ ID NO:7 is the amino acid sequence of Ac-AMP1.

SEQ ID NO:8 is the amino acid sequence of Ac-AMP2.

SEQ ID NO:9 is the nucleotide sequence of Ac-AMP2.

SEQ ID NO:10 is the amino acid sequence of Ac-AMP2 deduced from SED ID NO:9.

SEQ ID NO:11 is the amino acid sequence of Ca-AMP1.

SEQ ID NO:12 is one possible predicted DNA sequence for the Ca-AMP1 gene.

SEQ ID NO:13 is the amino acid sequence of Bm-AMP1.

SEQ ID NO:14 is one possible predicted DNA sequence for the Bm-AMP1 gene.

SEQ ID NO:15 is the amino acid sequence of Rs-AFP1.

SEQ ID NO:16 is the amino acid sequence of Rs-AFP2.

SEQ ID NO:17 is the amino acid sequence of Br-AFP1.

SEQ ID NO:18 is the amino acid sequence of Br-AFP2.

SEQ ID NO:19 is the amino acid sequence of

Bn-AFP1.

SEQ ID NO:20 is the amino acid sequence of Bn-AFP2.

SEQ ID NO:21 is the amino acid sequence of Sa-AFP1.

SEQ ID NO:22 is the amino acid sequence of Sa-AFP2.

SEQ ID NO:23 is the amino acid sequence of At-AFP1.

SEQ ID NO:24 is the amino acid sequence of Dm-AMP1.

SEQ ID NO:25 is the amino acid sequence of Dm-AMP2.

SEQ ID NO:26 is the amino acid sequence of Cb-AMP1.

SEQ ID NO:27 is the amino acid sequence of Cb-AMP2.

SEQ ID NO:28 is the amino acid sequence of Lc-AFP.

SEQ ID NO:29 is the amino acid sequence of Ct-AMP1.

SEQ ID NO:30 is the amino acid sequence of Rs-nsLTP.

SEQ ID NO:31 is one possible predicted DNA sequence for the Dm-AMP1 gene.

SEQ ID NO:32 is one possible predicted DNA sequence for the Dm-AMP2 gene.

SEQ ID NO:33 is one possible predicted DNA sequence for the Cb-AMP1 gene.

SEQ ID NO:34 is one possible predicted DNA sequence for the Cb-AMP2 gene.

SEQ ID NO:35 is one possible predicted DNA sequence for the Lc-AFP gene.

SEQ ID NO:36 is one possible predicted DNA sequence for the Ct-AMP1 gene.

SEQ ID NO:37 is the full length cDNA sequence

of Rs-AFP1.

SEQ ID NO:38 is the amino acid sequence of Rs-AFP1 deduced from SEQ ID NO:37.

SEQ ID NO:39 is the truncated cDNA sequence of Rs-AFP2.

SEQ ID NO:40 is the amino acid sequence of Rs-AFP2 deduced from SEQ ID NO:39.

SEQ ID NO:41 is the full length DNA sequence of PCR assisted site directed mutagenesis of Rs-AFP2.

SEQ ID NO:42 is the amino acid sequence of Rs-AFP2 deduced from SEQ ID NO:41.

EXAMPLE 1

Expression of Raphanus sativus Antifungal Protein 2 (Rs-AFP2) by the endophyte Clavibacter xyli subsp. cynodontis (Cxc).

The Rs-AFP2 protein is expressed in a system analogous to that which is known to express the Bacillus thuringiensis endotoxin. An oligonucleotide sequence coding for the antifungal protein Rs-AFP2 is prepared using Cxc-compatible codons. This oligonucleotide sequence comprises appropriate restriction sites to enable it to be exchanged with the Bacillus thuringiensis endotoxin gene sequence present in the INCIDE Cxc bacterium.

Southern analysis is used to check that Cxc is transformed with the Rs-AFP2 gene. If the result is positive, the bacterium is cultured to determine whether it is capable of expressing Rs-AFP2 protein in vitro. Western analysis and antifungal assays are carried out on the fermentation products to

determine whether the protein is produced in the correctly folded form as found in the native plant. It is known that the protein loses antifungal activity when it is reduced and hence unfolded.

EXAMPLE 2

Protection of rice plants using Rs-AFP2-producing Cxc as an antifungal agent.

Cultures of Cxc which are capable of expressing Rs-AFP2 protein are used to treat rice plants by a soil drench or seed treatment method.

The rice plants are challenged with rice blast, Pyricularia oryzae and assessed for increased resistance to the pathogen over non-Cxc-infected plants. Rs-AFP2 is known to be active against P oryzae in in vitro tests.

SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANT: ZENECA, Limited

(ii) TITLE OF INVENTION: ANITMICROBIAL-PROTEIN-PRODUCING ENDOSYMBIOTIC MICRO-ORGANISMS

(iii) NUMBER OF SEQUENCES: 42

(iv) CORRESPONDENCE ADDRESS:
(A) ADDRESSEE: ICI GROUP PATENTS SERVICES DEPT
(B) STREET: PO BOX 6, SHIRE PARK, BESSEMER ROAD,
(C) CITY: WELWYN GARDEN CITY
(D) STATE: HERTFORDSHIRE
(E) COUNTRY: UNITED KINGDOM
(F) ZIP: AL7 1HD

(v) COMPUTER READABLE FORM:
(A) MEDIUM TYPE: Floppy disk
(B) COMPUTER: IBM PC compatible
(C) OPERATING SYSTEM: PC-DOS/MS-DOS
(D) SOFTWARE: PatentIn Release #1.0, Version #1.25

(vi) CURRENT APPLICATION DATA:
(A) APPLICATION NUMBER:
(B) FILING DATE:
(C) CLASSIFICATION:

(vii) PRIOR APPLICATION DATA:
(A) APPLICATION NUMBER: GB 9300281.4
(B) FILING DATE: 08-JAN-1993

(viii) ATTORNEY/AGENT INFORMATION:
(A) NAME: ROBERTS, TIMOTHY W

(ix) TELECOMMUNICATION INFORMATION:
(A) TELEPHONE: 44 707 323400
(B) TELEFAX: 44 707 337454
(C) TELEX: 94028500 ICIC G

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 37 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

Gln Cys Ile Gly Asn Gly Gly Arg Cys Asn Glu Asn Val Gly Pro Pro
1 5 10 15

Tyr Cys Cys Ser Gly Phe Cys Leu Arg Gln Pro Gly Gln Gly Tyr Gly
20 25 30

Tyr Cys Lys Asn Arg
35

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 36 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

Cys Ile Gly Asn Gly Gly Arg Cys Asn Glu Asn Val Gly Pro Pro Tyr
1 5 10 15

Cys Cys Ser Gly Phe Cys Leu Arg Gln Pro Asn Gln Gly Tyr Gly Val
20 25 30

Cys Arg Asn Arg
35

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 360 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

CTTCCCGTTG CCTTCCTCAA ATTCGCTATT GTGTTGATTC TCTTCATTGC CATGTCCGCA 60

ATGATAGAAG CACAATGCAT AGGAAATGGA GGAAGATGTA ACGAGAACGT GGGGCCACCA 120

TACTGCTGCT CCGGTTTCTG CCTCCGTCAA CCTGGACAAG GTTATGGATA TTGTAAGAAC 180

CGCTGAGCAA GAGCATGAAA GCAAGGCCAA TGTGTGGTCT ACTAATTTAG CCTCAAATGT 240

TATTTATTTG CATGTCTTGT GTTTCTTAAT TACCTTCTTT GTGTCTAAGA AGGTATAGAT 300

CAATAGTTTC TACTTTACTA CTATGAATAA GAGGCTTTGA TTTGGTTTAA AAAAAAAAAA 360

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 61 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

Leu Pro Val Ala Phe Leu Lys Phe Ala Ile Val Leu Ile Leu Phe Ile
1 5 10 15

Ala Met Ser Ala Met Ile Glu Ala Gln Cys Ile Gly Asn Gly Gly Arg
20 25 30

Cys Asn Glu Asn Val Gly Pro Pro Tyr Cys Cys Ser Gly Phe Cys Leu
35 40 45

Arg Gln Pro Gly Gln Gly Tyr Gly Tyr Cys Lys Asn Arg
50 55 60

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 433 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

ATATCATTCA AATATACTAA ACTAATTATA AAAAATGGCT AAGGTTCCAA TTGCCTTTCT 60

CAAATTCGTC ATCGTGTTGA TTCTCTTCAT TGCCATGTCA GGCATGATAG AAGCATGCAT 120

AGGAAATGGA GGAAGATGTA ACGAGAACGT GGGCCCACCA TACTGCTGTT CGGGTTTCTG 180

CCTCCGTCAA CCTAACCAAG GTTACGGTGT TTGCAGGAAC CGCTAATAAG CAAAGCCCAA 240

AGTGTGGGTC ACAAAATAGT AGAGTTTAGC CTCAAATGTG GTTTATATAT GTAACAATCT 300

TATATGTGTT TCTCTTGTGT TTCTTAATTA CCTTCTTTGT GTCTAAGAAG GTATGGATAA 360

ATAGTTTGTA CTTTACTATT ATGGTTTTTT CTTATATCAA TAAGAGGCTT TAATTAAAAA 420

AAAAAAAAAA AAA 433

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 63 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

Met Ala Lys Val Pro Ile Ala Phe Leu Lys Phe Val Ile Val Leu Ile
1 5 10 15

Leu Phe Ile Ala Met Ser Gly Met Ile Glu Ala Cys Ile Gly Asn Gly
20 25 30

Gly Arg Cys Asn Glu Asn Val Gly Pro Pro Tyr Cys Cys Ser Gly Phe
35 40 45

Cys Leu Arg Gln Pro Asn Gln Gly Tyr Gly Val Cys Arg Asn Arg
50 55 60

(2) INFORMATION FOR SEQ ID NO:7:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 29 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

Val Gly Glu Cys Val Arg Gly Arg Cys Pro Ser Gly Met Cys Cys Ser

1 5 10 15

Gln Phe Gly Tyr Cys Gly Lys Gly Pro Lys Tyr Cys Gly
20 25

(2) INFORMATION FOR SEQ ID NO:8:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 30 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

Val Gly Glu Cys Val Arg Gly Arg Cys Pro Ser Gly Met Cys Cys Ser
1 5 10 15

Gln Phe Gly Tyr Cys Gly Lys Gly Pro Lys Tyr Cys Gly Arg
20 25 30

(2) INFORMATION FOR SEQ ID NO:9:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 590 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

CAAAAAAAAA AAATAAAGTC AAGAGTATTA ATTAGGTGAG AAAAAATGGT GAACATGAAG 60

TGTGTTGCAT TGATAGTTAT AGTTATGATG GCGTTTATGA TGGTGGATCC ATCAATGGGA 120

GTGGGAGAAT GTGTGAGAGG ACGTTGCCCA AGTGGGATGT GTTGCAGTCA GTTTGGGTAC 180

TGTGGTAAAG GCCCAAAGTA CTGTGGCCGT GCCAGTACTA CTGTGGATCA CCAAGCTGAT 240

GTTGCTGCCA CCAAAACTGC CAAGAATCCT ACCGATGCTA AACTTGCTGG TGCTGGTAGT 300

CCATGAAAGT AGTAGCTAGC TAGGTTCACG TTGGATTACC AAGCCGTGCC AGTACTACTG 360

TGGCCGTGCC AGTACTAATG TTCTCTTATA TGTCTGAAAT AAGCTCCTAT ATAAATACTA 420

GTATCTTGAT GTAATGGAGT ATTTTCATTT TGTTTTTATT TGAGTTATGA TCGTGACTTC 480

CTTGTGTTGG TTTAACTTGT ATATTGTAAT GCATCTTAAA TGCTGTCTCA AATAATTTGA 540

TGTATTAAAC ACTTGTTTTG TTTTTAATAC ATACTAAGTG CTGTAAATTC 590

(2) INFORMATION FOR SEQ ID NO:10:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 86 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

Met Val Asn Met Lys Cys Val Ala Leu Ile Val Ile Val Met Met Ala
1 5 10 15

Phe Met Met Val Asp Pro Ser Met Gly Val Gly Glu Cys Val Arg Gly
20 25 30

Arg Cys Pro Ser Gly Met Cys Cys Ser Gln Phe Gly Tyr Cys Gly Lys
35 40 45

Gly Pro Lys Tyr Cys Gly Arg Ala Ser Thr Thr Val Asp His Gln Ala
50 55 60

Asp Val Ala Ala Thr Lys Thr Ala Lys Asn Pro Thr Asp Ala Lys Leu
65 70 75 80

Ala Gly Ala Gly Ser Pro
85

(2) INFORMATION FOR SEQ ID NO:11:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 42 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

Gln Glu Gln Cys Gly Asn Gln Ala Gly Gly Arg Ala Cys Ala Asn Arg
1 5 10 15

Leu Cys Cys Ser Gln Tyr Gly Tyr Cys Gly Ser Thr Arg Ala Tyr Cys
20 25 30

Gly Val Gly Cys Gln Ser Asn Cys Gly Arg
35 40

(2) INFORMATION FOR SEQ ID NO:12:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 126 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

CAAGAGCAAT GCGGAAACCA AGCTGGAGGA AGAGCTTGCG CTAACAGACT TTGCTGCTCT 60

CAATACGGAT ACTGCGGATC TACTAGAGCT TACTGCGGAG TTGGATGCCA ATCTAACTGC 120

GGAAGA 126

(2) INFORMATION FOR SEQ ID NO:13:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 37 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 15
(D) OTHER INFORMATION: /note= "Xaa at position 15 may be R or H"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 29
(D) OTHER INFORMATION: /note= "Xaa at position 29 may be S or N"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

Cys Ser Ser His Asn Pro Cys Pro Arg His Gln Cys Cys Ser Xaa Tyr
1 5 10 15

Gly Tyr Cys Gly Leu Gly Ser Asp Tyr Cys Gly Leu Xaa Cys Arg Gly
20 25 30

Gly Pro Cys Asp Arg
35

(2) INFORMATION FOR SEQ ID NO:14:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 111 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

TGCTCTTCTC ACAACCCGTG CCCGAGACAC CAATGCTGCT CTAAGTACGG ATACTGCGGA 60

CTTGGATCTG ACTACTGCGG ACTTGGATGC AGAGGAGGAC CGTGCGACAG A 111

(2) INFORMATION FOR SEQ ID NO:15:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 44 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

Gln Lys Leu Cys Glu Arg Pro Ser Gly Thr Trp Ser Gly Val Cys Gly
1 5 10 15

Asn Asn Asn Ala Cys Lys Asn Gln Cys Ile Asn Leu Glu Lys Ala Arg
20 25 30

His Gly Ser Cys Asn Tyr Val Phe Pro Ala His Lys
35 40

(2) INFORMATION FOR SEQ ID NO:16:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 36 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

Gln Lys Leu Cys Gln Arg Pro Ser Gly Thr Trp Ser Gly Val Cys Gly
1 5 10 15

Asn Asn Asn Ala Cys Lys Asn Gln Cys Ile Arg Leu Glu Lys Ala Arg
20 25 30

His Gly Ser Cys
35

(2) INFORMATION FOR SEQ ID NO:17:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 27 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:

Gln Lys Leu Cys Glu Arg Pro Ser Gly Thr Trp Ser Gly Val Cys Gly
1 5 10 15

Asn Asn Asn Ala Cys Lys Asn Gln Cys Ile Asn
20 25

(2) INFORMATION FOR SEQ ID NO:18:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 27 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:

Gln Lys Leu Cys Glu Arg Pro Ser Gly Thr Xaa Ser Gly Val Cys Gly
1 5 10 15

Asn Asn Asn Ala Cys Lys Asn Gln Cys Ile Arg
20 25

(2) INFORMATION FOR SEQ ID NO:19:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 30 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:19:

Gln Lys Leu Cys Glu Arg Pro Ser Gly Thr Trp Ser Gly Val Cys Gly
1 5 10 15

Asn Asn Asn Ala Cys Lys Asn Gln Cys Ile Asn Leu Glu Lys
20 25 30

(2) INFORMATION FOR SEQ ID NO:20:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 23 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:20:

Gln Lys Leu Cys Glu Arg Pro Ser Gly Thr Trp Ser Gly Val Cys Gly
1 5 10 15

Asn Asn Asn Ala Cys Lys Asn
20

(2) INFORMATION FOR SEQ ID NO:21:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 25 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:21:

Gln Lys Leu Cys Glu Arg Pro Ser Gly Thr Trp Ser Gly Val Cys Gly
1 5 10 15

Asn Asn Asn Ala Cys Lys Asn Gln Cys
20 25

(2) INFORMATION FOR SEQ ID NO:22:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 26 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:22:

Gln Lys Leu Cys Gln Arg Pro Ser Gly Thr Trp Ser Gly Val Cys Gly
1 5 10 15

Asn Asn Asn Ala Cys Arg Asn Gln Cys Ile
20 25

(2) INFORMATION FOR SEQ ID NO:23:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 27 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:23:

Gln Lys Leu Cys Glu Arg Pro Ser Gly Thr Trp Ser Gly Val Cys Gly
1 5 10 15

Asn Ser Asn Ala Cys Lys Asn Gln Cys Ile Asn
20 25

(2) INFORMATION FOR SEQ ID NO:24:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 50 amino acids
(B) TYPE: amino acid

(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:24:

Glu Leu Cys Glu Lys Ala Ser Lys Thr Trp Ser Gly Asn Cys Gly Asn
1 5 10 15

Thr Gly His Cys Asp Asn Gln Cys Lys Ser Trp Glu Gly Ala Ala His
20 25 30

Gly Ala Cys His Val Arg Asn Gly Lys His Met Cys Phe Cys Tyr Phe
35 40 45

Asn Cys
50

(2) INFORMATION FOR SEQ ID NO:25:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:25:

Glu Val Cys Glu Lys Ala Ser Lys Thr Trp Ser Gly Asn Cys Gly Asn
1 5 10 15

Thr Gly His Cys
20

(2) INFORMATION FOR SEQ ID NO:26:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 50 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:26:

Glu Leu Cys Glu Lys Ala Ser Lys Thr Trp Ser Gly Asn Cys Gly Asn
1 5 10 15

Thr Lys His Cys Asp Asp Gln Cys Lys Ser Trp Glu Gly Ala Ala His
20 25 30

Gly Ala Cys His Val Arg Asn Gly Lys His Met Cys Phe Cys Tyr Phe
35 40 45

Asn Cys
50

(2) INFORMATION FOR SEQ ID NO:27:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 50 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:27:

Glu Leu Cys Glu Lys Ala Ser Lys Thr Trp Ser Gly Asn Cys Gly Asn
1 5 10 15

Thr Lys His Cys Asp Asn Lys Cys Lys Ser Trp Glu Gly Ala Ala His
20 25 30

Gly Ala Cys His Val Arg Ser Gly Lys His Met Cys Phe Cys Tyr Phe
35 40 45

Asn Cys
50

(2) INFORMATION FOR SEQ ID NO:28:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 47 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:28:

Lys Thr Cys Glu Asn Leu Ser Gly Thr Phe Lys Gly Pro Cys Ile Pro

1 5 10 15

Asp Gly Asn Cys Asn Lys His Cys Lys Asn Asn Glu His Leu Leu Ser
20 25 30

Gly Arg Cys Arg Asp Asp Phe Xaa Cys Trp Cys Thr Arg Asn Cys
35 40 45

(2) INFORMATION FOR SEQ ID NO:29:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 49 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:29:

Asn Leu Cys Glu Arg Ala Ser Leu Thr Trp Thr Gly Asn Cys Gly Asn
1 5 10 15

Thr Gly His Cys Asp Thr Gln Cys Arg Asn Trp Glu Ser Ala Lys His
20 25 30

Gly Ala Cys His Lys Arg Gly Asn Trp Lys Cys Phe Cys Tyr Phe Asp
35 40 45

Cys

(2) INFORMATION FOR SEQ ID NO:30:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 43 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:30:

Ala Leu Ser Cys Gly Thr Val Asn Ser Asn Leu Ala Ala Cys Ile Gly
1 5 10 15

Tyr Leu Thr Gln Asn Ala Pro Leu Ala Arg Gly Cys Cys Thr Gly Val
20 25 30

Thr Asn Leu Asn Asn Met Ala Xaa Thr Thr Pro
35 40

(2) INFORMATION FOR SEQ ID NO:31:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 150 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:31:

GAGCTTTGCG AGAAGGCTTC TAAGACTTGG TCTGGAAACT GCGGAAACAC TGGACATTGC 60

GATAACCAAT GCAAGTCTTG GGAGGGAGCT GCTCATGGAG CTTGCCATGT TAGAAACGGA 120

AAGCATATGT GCTTCTGCTA CTTCAACTGC 150

(2) INFORMATION FOR SEQ ID NO:32:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 60 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:32:

GAGGTTTGCG AGAAGGCTTC TAAGACTTGG TCTGGAAACT GCGGAAACAC TGGACATTGC 60

(2) INFORMATION FOR SEQ ID NO:33:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 150 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:33:

GAGCTTTGCG AGAAGGCTTC TAAGACTTGG TCTGGAAACT GCGGAAACAC TAAGCATTGC 60

GATGATCAAT GCAAGTCTTG GGAGGGAGCT GCTCATGGAG CTTGCCATGT TAGAAACGGA 120

AAGCATATGT GCTTCTGCTA CTTCAACTGC 150

(2) INFORMATION FOR SEQ ID NO:34:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 150 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:34:

GAGCTTTGCG AGAAGGCTTC TAAGACTTGG TCTGGAAACT GCGGAAACAC TAAGCATTGC 60

GATAACAAGT GCAAGTCTTG GGAGGGAGCT GCTCATGGAG CTTGCCATGT TAGATCTGGA 120

AAGCATATGT GCTTCTGCTA CTTCAACTGC 150

(2) INFORMATION FOR SEQ ID NO:35:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 141 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:35:

AAGACTTGCG AGAACCTTTC TGGAACTTTC AAGGGACCAT GCATTCCAGA TGGAAACTGC 60

AACAAGCATT GCAAGAACAA CGAGCATCTT CTTTCTGGAA GATGCAGAGA TGATTTCNNN 120

TGCTGGTGCA CTAGAAACTG C 141

(2) INFORMATION FOR SEQ ID NO:36:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 147 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:36:

AACCTTTGCG AGAGAGCTTC TCTTACTTGG ACTGGAAACT GCGGAAACAC TGGACATTGC 60

GATACTCAAT GCAGAAACTG GGAGTCTGCT AAGCATGGAG CTTGCCATAA GAGAGGAAAC 120

TGGAAGTGCT TCTGCTACTT CGATTGC 147

(2) INFORMATION FOR SEQ ID NO:37:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 414 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:37:

GTTTTATTAG TGATCATGGC TAAGTTTGCG TCCATCATCG CACTTCTTTT TGCTGCTCTT 60

GTTCTTTTTG CTGCTTTCGA AGCACCAACA ATGGTGGAAG CACAGAAGTT GTGCGAAAGG 120

CCAAGTGGGA CATGGTCAGG AGTCTGTGGA AACAATAACG CATGCAAGAA TCAGTGCATT 180

AACCTTGAGA AAGCACGACA TGGATCTTGC AACTATGTCT TCCCAGCTCA CAAGTGTATC 240

TGCTACTTTC CTTGTTAATT TATCGCAAAC TCTTTGGTGA ATAGTTTTTA TGTAATTTAC 300

ACAAAATAAG TCAGTGTCAC TATCCATGAG TGATTTTAAG ACATGTACCA GATATGTTAT 360

GTTGGTTCGG TTATACAAAT AAAGTTTTAT TCACCAAAAA AAAAAAAAAA AAAA 414

(2) INFORMATION FOR SEQ ID NO:38:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 80 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:38:

Met Ala Lys Phe Ala Ser Ile Ile Ala Leu Leu Phe Ala Ala Leu Val
1 5 10 15

Leu Phe Ala Ala Phe Glu Ala Pro Thr Met Val Glu Ala Gln Lys Leu
20 25 30

Cys Glu Arg Pro Ser Gly Thr Trp Ser Gly Val Cys Gly Asn Asn Asn
35 40 45

Ala Cys Lys Asn Gln Cys Ile Asn Leu Glu Lys Ala Arg His Gly Ser
50 55 60

Cys Asn Tyr Val Phe Pro Ala His Lys Cys Ile Cys Tyr Phe Pro Cys
65 70 75 80

(2) INFORMATION FOR SEQ ID NO:39:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 284 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:39:

GGAAATAATA ACGCATGCAA GAATCAGTGC ATTCGACTTG AGAAAGCACG ACATGGGTCT 60

TGCAACTATG TCTTCCCAGC TCACAAGTGT ATCTGTTATT TCCCTTGTTA ATTCCATAAA 120

CTCTTCGGTG GTTAATAGTG TGCGCATATT ACATATAATT AATAAGTTTG TGTCACTATT 180

TATTAGTGAC TTTATGACAT GTGCCAGGTA TGTTTATGTT GGGTTGGTTG TAATATAAAA 240

AAGTTCACGG ATAATAAGAT GATAAGCTCA CGTCGCCAAA AAAA 284

(2) INFORMATION FOR SEQ ID NO:40:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 36 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:40:

Gly Asn Asn Asn Ala Cys Lys Asn Gln Cys Ile Arg Leu Glu Lys Ala
1 5 10 15

Arg His Gly Ser Cys Asn Tyr Val Phe Pro Ala His Lys Cys Ile Cys
20 25 30

Tyr Phe Pro Cys
35

(2) INFORMATION FOR SEQ ID NO:41:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 288 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:41:

CCCCGGGCTG CAGGAATTCG CGGCCGCGTT TTATTAGTGA TCATGGCTAA GTTTGCGTCC 60

ATCATCGCAC TTCTTTTTGC TGCTCTTGTT CTTTTTGCTG CTTTCGAAGC ACCAACAATG 120

GTGGAAGCAC AGAAGTTGTG CCAAAGGCCA AGTGGGACAT GGTCAGGAGT CTGTGGAAAC 180

AATAACGCAT GCAAGAATCA GTGCATTAGA CTTGAGAAAG CACGACATGG ATCTTGCAAC 240

TATGTCTTCC CAGCTCACAA GTGTATCTGC TACTTTCCTT GTTAATAG 288

(2) INFORMATION FOR SEQ ID NO:42:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 80 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:42:

Met Ala Lys Phe Ala Ser Ile Ile Ala Leu Leu Phe Ala Ala Leu Val
1 5 10 15

Leu Phe Ala Ala Phe Glu Ala Pro Thr Met Val Glu Ala Gln Lys Leu
20 25 30

Cys Gln Arg Pro Ser Gly Thr Trp Ser Gly Val Cys Gly Asn Asn Asn

35 40 45

Ala Cys Lys Asn Gln Cys Ile Arg Leu Glu Lys Ala Arg His Gly Ser
50 55 60

Cys Asn Tyr Val Phe Pro Ala His Lys Cys Ile Cys Tyr Phe Pro Cys
65 70 75 80

We claim:

1. A method of producing an antimicrobial-protein-producing micro-organism capable of entering into an endosymbiotic relationship with a plant host comprising the combination of genetic material encoding a plant-derived antimicrobial protein with an endophyte.

2. A method according to claim 1 in which the plant-derived antimicrobial protein is selected from the protein group consisting of Mj-AMP1, Mj-AMP2, Ac-AMP1, Ac-AMP2, Ca-AMP1, Bm-AMP1, Rs-AFP1, Rs-AFP2, Br-AFP1, Br-AFP2, Bn-AFP1, Bn-AFP2, Sa-AFP1, Sa-AFP2, At-AFP1, Dm-AMP1, Dm-AMP2, Cb-AMP1, Cb-AMP2, Lc-AFP, Ct-AMP1, Ct-AMP2 and Rs-nsLTP.

3. A method according to claim 1 in which the endopyte is Clavibacter xyli subsp. cynodontis.

4. An antimicrobial-protein-producing micro-organism produced by the method according to claim 1.

5. A method for protecting a plant host from disease comprising treating the plant host with the antimicrobial-protein-producing micro-organism according to claim 4.

6. A plant or seed treated with an antimicrobial-protein-producing micro-organism according to claim 4.